

GRADIENT DYNAMICAL SYSTEMS ON OPEN SURFACES AND CRITICAL POINTS OF GREEN'S FUNCTIONS

ALBERTO ENCISO AND DANIEL PERALTA-SALAS

ABSTRACT. We study the dynamics of the vector field on an open surface given by the gradient of a Green's function. This dynamical approach enables us to show that this field induces an invariant decomposition of the surface as the union of a disk and a 1-skeleton that encodes the topology of the surface. We analyze the structure of this 1-skeleton, thereby obtaining, in particular, a topological upper bound for the number of critical points a Green's function can have. Connections between the dynamical properties of the gradient field and the conformal structure of the surface are also discussed.

1. INTRODUCTION

Let M be a noncompact surface without boundary of class C^∞ , endowed with a smooth complete Riemannian metric g . We denote by

$$\mathcal{G} : (M \times M) \setminus \text{diag}(M \times M) \rightarrow \mathbb{R}$$

a Green's function of M , which is defined as a symmetric function (i.e., $\mathcal{G}(x, y) = \mathcal{G}(y, x)$) that satisfies the equation

$$(1.1) \quad \Delta_g \mathcal{G}(\cdot, y) = -\delta_y$$

for each $y \in M$. That is to say, if one considers the action of the Laplace–Beltrami operator of the manifold, Δ_g , on the Green's function $\mathcal{G}(x, y)$ (with respect to the first variable x), one obtains a Dirac measure supported at the point y .

Our goal in this paper is to analyze the dynamical properties of the gradient of the Green's function. For this, we will find it notationally convenient to fix a point $y \in M$, once and for all, and use the notation $G := \mathcal{G}(\cdot, y)$ for the Green's function with pole y , which is smooth and harmonic in $M \setminus \{y\}$. Therefore, the gradient field we will study in this paper will be $\nabla_g G$.

The study of Green's functions is a central topic in Riemannian geometry and geometric analysis. Hence, there is a vast related literature covering, among many other aspects, the existence of positive Green's functions [3, 11], upper and lower bounds, gradient estimates and asymptotics [13, 4], and the connection between Green's functions and the heat kernel [12, 10].

What is somewhat surprising is that the dynamical properties of the gradient field $\nabla_g G$ remain virtually unexplored, with the exception of the classical work of Brelot and Choquet [1]. Of course, a key issue in the study of this vector field is the analysis of the critical set of the Green's function. This question is of considerable interest by itself, and deeply related with other problems, set in significantly easier contexts, that date back to the 1950s (see e.g. [20, 17, 16, 19] and references therein). Indeed, some of these articles were motivated in part by

the fact that in Euclidean space the Green's function arises as the electric potential of a charged particle, so that its critical points correspond to equilibria and the trajectories of its gradient field are the force lines studied by Faraday and Maxwell in the XIX century (nontrivial contributions to this problem were made in the recent paper [7]). As a side remark, let us point out that another elliptic PDE (very different from (1.1)) in which the analysis of the critical points of its solutions has recently attracted considerable attention can be found in [14].

One reason why the study of the dynamical properties of the gradient field $\nabla_g G$ (and in particular of the critical set of G) is hard to tackle, from the point of view of geometric analysis, is that the estimates for Green's functions obtained through PDE methods are not sufficiently fine to elucidate whether the gradient of G vanishes in a certain region. Moreover, the noncompactness of the underlying surface introduces complications related to the behavior of the Green's function at infinity.

In this paper we will show how these difficulties can be overcome by exploiting the conformal properties of the surface and resorting to techniques of gradient dynamical systems. Our approach will lead to a topological upper bound for the number of critical points of the Green's function and a complete description of the local and global dynamics of the gradient field $\nabla_g G$.

To some extent, the core of this paper is the remarkable heuristic principle we will now state, which links the dynamics of the gradient field $\nabla_g G$, defined in terms of solutions to an elliptic PDE, with the topology of the underlying surface. It must be stressed that this principle will be promoted to a rigorous statement (after introducing the necessary tools and notation) in Corollary 4.6 and Theorem 5.4:

Heuristic principle. *Suppose that the surface M is of finite type. Then M can be decomposed as the union of a disk D and a (possibly disconnected) noncompact graph \mathcal{F} , both of which are invariant under the local flow of the gradient field $\nabla_g G$. The disk consists of the pole y and the points of M whose ω -limit is y . The graph \mathcal{F} consists of the critical points of G , their stable components, and certain trajectories of $\nabla_g G$ that escape to infinity. When suitably compactified, \mathcal{F} is a connected graph that encodes the topology of the surface, the rank of the first homology group of \mathcal{F} being twice the genus of M .*

The characterization of the set \mathcal{F} that will emerge from the rigorous version of this heuristic principle yields, as a nontrivial application, the following topological upper bound for the number of critical points of the Green's function:

Theorem 1.1. *Suppose that the surface M is of finite type, that is, its fundamental group has finite rank. Then the number of critical points of any Green's function G on M is not larger than twice the genus of M , ν , plus the number of ends, λ , minus 1:*

$$\# \text{critical points} \leq 2\nu + \lambda - 1.$$

If this upper bound is attained then G is a Morse function.

In fact, the analysis of the set \mathcal{F} does not only yield this topological upper bound, but a more refined bound that exploits the conformal structure of the surface (see Theorem 5.1). This is particularly interesting because, as we shall see, it establishes some links between the conformal geometry of the surface and the portrait of the

gradient field $\nabla_g G$. It should be stressed that an analogous result does not hold for higher-dimensional Riemannian manifolds, as shown in [6].

A different proof of the estimates for the number of critical points of G , relying on methods from elliptic PDEs, was recently given by the authors in [6]. However, the dynamical approach taken in the present paper provides a very satisfactory picture of the invariant sets connecting the different critical points of G and the dynamics of the field $\nabla_g G$, which cannot be obtained with the PDE techniques used in [6].

The article is organized as follows. In Section 2 we will present some basic facts regarding Green's functions on surfaces, including their obtention through an exhaustion procedure, their behavior at infinity and their connection with the conformal structure of the surface. In Section 3 we describe the dynamics of the field $\nabla_g G$ in a neighborhood of the pole y or a critical point. In Section 4 we introduce a convenient compactification of the surface and establish some key properties of the sets D and \mathcal{F} introduced in the Heuristic Principle above (and of some compactifications thereof). The structure of the set \mathcal{F} and its compactification is characterized in Section 5, which allows us to prove the upper bound for the number of critical points of the Green's function. To conclude, in Section 6 we discuss the connection between the dynamics of the field $\nabla_g G$ and the conformal geometry of the underlying surface and make some comments regarding surfaces of infinite topological type.

2. GREEN'S FUNCTIONS ON SURFACES

In this section we will recall what a Green's function is and how to obtain them using an exhaustion procedure, placing a special emphasis on conditions ensuring that the Green's function is "well behaved" at infinity, which is a key ingredient in the analysis of the dynamical properties of the vector field $\nabla_g G$. We will also discuss how to exploit conformal isometries to classify the possible behavior of the Green's function at the ends of the surface. Throughout this paper, the surfaces will be of class C^∞ and of finite topological type (that is, with finitely generated fundamental group) unless stated otherwise.

The reason why it is crucial to make assumptions on the behavior of the Green's function at infinity can be readily seen even in the simplest case: the Euclidean plane \mathbb{R}^2 . Indeed, if we let h be any harmonic function, it is clear that any function of the form

$$\mathcal{G}(x, y) = -\frac{1}{2\pi} \log |x - y| + h(x) + h(y)$$

is symmetric and satisfies the Green's function equation (1.1). The standard way of deciding which of these Green's functions should be "admissible" is to demand that the Green's function be obtained as a limit of Dirichlet Green's functions associated with an exhaustion of the plane by compact subsets (more details on this point will be given below). The only Green's function arising from such an exhaustion procedure would be the usual one,

$$(2.1) \quad \mathcal{G}(x, y) = -\frac{1}{2\pi} \log |x - y|,$$

which is the relevant Green's function for all geometric or analytical considerations.

In a general Riemannian surface (M, g) , we will assume that the Green's function G we consider shares the following two properties with the above Green's function (2.1) of the Euclidean plane. The first assumption is a weak monotonicity property for the Green's function on circles. The second assumption says that, when the surface admits a *minimal (positive) Green's function* (as is the case of the hyperbolic plane, although not of the Euclidean one), we should always consider this Green's function, for it plays a very special role in analysis and geometry. Recall that a positive Green's function is minimal when it is pointwise smaller than any other positive Green's function. Observe that, if the surface (M, g) does not admit a minimal Green's function, then the infimum of any Green's function G over the surface is $-\infty$.

Assumption 1 (Monotonicity). The Green's function G is nondecreasing in the sense that

$$(2.2) \quad \sup_{M \setminus B_g(y, r)} G = \max_{\partial B_g(y, r)} G \quad \text{for all } r > 0,$$

where $B_g(y, r)$ denotes the geodesic disk in M centered at the pole y of radius r .

Assumption 2 (Minimality). G is the minimal Green's function whenever the Riemannian surface admits a positive Green's function.

It should be stressed that on any Riemannian surface (M, g) there are Green's functions that satisfy these assumptions. Indeed, the way one shows there always exist Green's functions on any surface is by taking a suitable limit of the Dirichlet Green's functions of an exhaustion of the surface by bounded domains. Since the Green's functions one obtains in this fashion actually satisfy the above assumptions, *throughout this paper we will assume that the Green's function satisfies Assumptions 1 and 2*. For the benefit of the reader, we will next review this construction of Green's functions through an exhaustion by compact sets, which was introduced in the context of general Riemannian manifolds by Li and Tam [11]. To simplify the exposition, we will consider the case of Green's functions $G(x)$ with a fixed pole y , but actually the procedure automatically yields the symmetric function $\mathcal{G}(x, y)$.

Consider an exhaustion $\Omega_1 \subset \Omega_2 \subset \dots$ of the surface M by bounded domains. We can assume without loss of generality that the pole y belongs to the first domain Ω_1 . The idea now is to impose Dirichlet boundary conditions on each bounded domain Ω_j and consider the corresponding Green's function $G_j : \Omega_j \setminus \{y\} \rightarrow \mathbb{R}$, which satisfies the equation

$$\Delta_g G_j = -\delta_y \quad \text{in } \Omega_j, \quad G_j = 0 \quad \text{on } \partial\Omega_j.$$

One would be tempted to define the Green's function G as the limit of G_j as $j \rightarrow \infty$. However, this limit does not exist in general. What can be proved (see e.g. [11]) is that, for any choice of the domains Ω_j , one can take a sequence of nonnegative real numbers $(a_j)_{j=1}^\infty$ such that $G_j - a_j$ converges uniformly on compact sets of $M \setminus \{y\}$ to a Green's function G with pole y . Generally, the Green's functions obtained through this procedure are non-unique, but they satisfy the monotonicity assumption (2.2) and, when the surface admits a positive Green's function, this construction always yields the minimal one.

Incidentally, it is worth pointing out that Green's functions do not exist on closed surfaces, which is the reason why we just consider noncompact surfaces in

this paper. To see why, it is enough to suppose that there is a solution of Eq. (1.1) in a closed surface M and, with a slight abuse of notation, integrate both sides of the equation over the whole surface and integrate by parts, which would yield the contradiction

$$0 = \int_M \Delta_g G = - \int_M \delta_p = -1.$$

Let us now pass to describe how one can utilize conformal isometries to analyze the behavior of the Green's function at each end. Recall that two Riemannian surfaces (M, g) and (\mathcal{M}, g_0) are *conformally isometric* if there is a diffeomorphism $\Phi : \mathcal{M} \rightarrow M$ and a smooth positive function f on \mathcal{M} such that $\Phi^*g = fg_0$. An important property of the Laplace equation on surfaces is its *conformal invariance*, that is, if $G(x)$ satisfies the equation

$$\Delta_g G = -\delta_y$$

on the surface M for some point y , and another surface \mathcal{M} is conformally isometric to M through a diffeomorphism $\Phi : \mathcal{M} \rightarrow M$, then

$$\overline{G}(x) := G(\Phi(x))$$

is a Green's function of \mathcal{M} with pole $\bar{y} := \Phi^{-1}(y)$:

$$(2.3) \quad \Delta_{g_0} \overline{G} = -\delta_{\bar{y}}.$$

Furthermore, the corresponding gradient fields are orbitally conjugated through the relation

$$(2.4) \quad \nabla_{g_0} \overline{G} = f \Phi^*(\nabla_g G).$$

A key result in the conformal geometry of surfaces, which will be of great use in this paper, is the uniformization theorem:

Theorem 2.1 (Uniformization). *There is a compact surface Σ of genus ν with a metric of constant curvature g_0 , a certain number $\lambda_1 \geq 0$ of isolated points $p_i \in \Sigma$ and another number $\lambda_2 \geq 0$ of closed disks $D_i \subset \Sigma$ with smooth boundary such that the Riemannian surface M is conformally isometric to (\mathcal{M}, g_0) , with*

$$(2.5) \quad \mathcal{M} := \Sigma \setminus \left(\bigcup_{i=1}^{\lambda_1} \{p_i\} \cup \bigcup_{j=1}^{\lambda_2} D_j \right).$$

As is customary, we will call the deleted points $\{p_1, \dots, p_{\lambda_1}\}$ the *parabolic ends* of the surface M , while the deleted disks $\{D_1, \dots, D_{\lambda_2}\}$ are its *hyperbolic ends*. Furthermore, we will say that the parabolic end p_i is a *removable singularity* if the function \overline{G} can be extended so as to satisfy the equation

$$\Delta_{g_0} \overline{G} = 0$$

in a neighborhood of p_i in Σ . It should be noticed that an end being parabolic or hyperbolic is a geometric property of the surface, related to its conformal structure. However, whether a parabolic end is removable or not is not a geometric issue, as it depends not only on the surface M but also on the Green's function we consider.

In the following two propositions we will relate the existence of parabolic and hyperbolic ends with the behavior of the function \overline{G} at each deleted point or disk:

Proposition 2.2. *If all the ends of the surface (M, g) are parabolic, the surface does not admit any positive Green's functions. Moreover, at each end p_i we have that either*

$$(2.6) \quad \lim_{x \rightarrow p_i} \bar{G}(x) = -\infty$$

or p_i is a removable singularity. There is at least one point p_i where the condition (2.6) is satisfied.

Proof. When the number of hyperbolic ends λ_2 is 0, it follows from Eq. (2.3) that the function \bar{G} satisfies the equation $\Delta_{g_0} \bar{G} = 0$ everywhere in Σ but at the pole \bar{y} and the isolated points p_i . Furthermore, by Assumption 1, \bar{G} is upper bounded at each point p_i . If it is also lower bounded, it is standard that p_i is a removable singularity [8] and $\Delta_{g_0} \bar{G} = 0$ in a neighborhood of p_i . If \bar{G} is not lower bounded, p_i is an isolated singularity of \bar{G} , and the fact that \bar{G} is upper bounded readily implies that $\Delta_{g_0} \bar{G} = c_i \delta_{p_i}$ in a neighborhood of p_i for some non-negative constant c_i . Hence

$$(2.7) \quad \Delta_{g_0} \bar{G} = -\delta_{\bar{y}} + \sum_{i=1}^{\lambda_1} c_i \delta_{p_i}$$

in the closed manifold Σ , so integrating this equation over Σ and using that $\int_{\Sigma} \Delta_{g_0} \bar{G} = 0$ we infer that $\sum_i c_i = 1$. Hence c_i is positive for some i and, in view of the asymptotic behavior of any Green's function at a pole, it stems that \bar{G} tends to $-\infty$ at the corresponding point p_i . \square

Proposition 2.3. *If the surface (M, g) has at least one hyperbolic end, there is a minimal positive Green's function G . The corresponding function \bar{G} tends to zero at the boundary of each disk D_i and all the parabolic ends p_i are removable singularities.*

Proof. When the number of hyperbolic ends is $\lambda_2 \geq 1$, it is standard (for example, due to the existence of nonconstant positive harmonic functions in \mathcal{M} [11]) that the surface admits a positive Green's function. Therefore, Assumption 2 ensures that G is the minimal Green's function of the surface, which corresponds to the unique solution \bar{G} of the boundary problem

$$\Delta_{g_0} \bar{G} = -\delta_{\bar{y}} \quad \text{in } \Sigma \setminus \bigcup_{j=1}^{\lambda_2} D_j, \quad \bar{G} = 0 \quad \text{on } \partial D_j \text{ for all } j.$$

In particular, if there are any parabolic ends, they are all removable singularities. \square

Because of these propositions, one can extend the function \bar{G} and the gradient field $\nabla_{g_0} \bar{G}$ to all the removable singularities of the surface. We will find it occasionally convenient to consider this extension, which we will not distinguish notationally (it will be clear from the context). To conclude this section, we will present two examples that illustrate the issue of uniqueness and non-uniqueness of Green's functions on surfaces.

Example 2.4. Consider the plane \mathbb{R}^2 with its Euclidean metric. It has a parabolic end, so that it is conformally isometric to the round sphere (\mathbb{S}^2, g_0) minus a point p via a diffeomorphism $\Phi : \mathbb{S}^2 \setminus \{p\} \rightarrow \mathbb{R}^2$.

The standard Green's function with pole y ,

$$(2.8) \quad G(x) := -\frac{1}{2\pi} \log |x - y|,$$

obviously satisfies Assumption 1. Moreover, it is the only Green's function satisfying this assumption, even though the plane does not admit any positive Green's functions. To see this, take the function \bar{G} corresponding to any Green's function satisfying Assumption 1. Since the assumption is satisfied, Eq. (2.7) in the proof of Proposition 2.2 shows that \bar{G} must satisfy the equation

$$\Delta_{g_0} \bar{G} = -\delta_{\bar{y}} + \delta_p$$

in the whole \mathbb{S}^2 . Hence \bar{G} is uniquely determined, so that the Green's function must be given by (2.8).

Example 2.5. Let us consider the flat cylinder $\mathbb{R} \times \mathbb{S}^1$ and natural coordinates (z, θ) . It is conformally equivalent to the round sphere (\mathbb{S}^2, g_0) minus two points $\{p_1, p_2\}$, so it does not admit a positive Green's function.

A Green's function with pole at (z_0, θ_0) is

$$G_1(z, \theta) := -\frac{1}{4\pi} \log [\cosh(z - z_0) - \cos(\theta - \theta_0)].$$

This Green's function satisfies Assumption 1. It is not the only Green's function on the flat cylinder with this property; e.g., one can check that

$$G_2(z, \theta) := -\frac{1}{4\pi} \log \left[\frac{1}{2} e^{2z} + \frac{1}{2} e^{2z_0} - e^{z+z_0} \cos(\theta - \theta_0) \right]$$

is another instance. These Green's functions are connected by the identity

$$G_2(z, \theta) = G_1(z, \theta) - \frac{z + z_0}{4\pi}.$$

Notice that G_1 tends to $-\infty$ at both ends (that is, as $z \rightarrow \pm\infty$) while G_2 tends to $-\infty$ as $z \rightarrow \infty$ but the end $z \rightarrow -\infty$ corresponds to a removable singularity of G_2 .

3. LOCAL DYNAMICAL PROPERTIES OF GREEN'S FUNCTIONS

In this section we will carry out a local study of the dynamics of the gradient of the Green's function in a neighborhood of the pole or a critical point. Since the fields $\nabla_g G$ and $\nabla_{g_0} \bar{G}$ are orbitally conjugated (cf. Eq. (2.4)), for convenience we will work with the latter gradient field instead.

The first proposition we will prove here asserts that, when multiplied by a suitable factor, the gradient vector field $\nabla_{g_0} \bar{G}$ can be smoothly linearized at the point \bar{y} , and that the corresponding normal form is a stable node. In particular, the trajectories approach the pole with a well-defined tangent.

Proposition 3.1. *There are C^∞ coordinates (x_1, x_2) , defined in a neighborhood U of the point \bar{y} in \mathcal{M} and centered at \bar{y} , and a smooth nonnegative function $\rho : U \rightarrow \mathbb{R}$ that only vanishes at the pole, such that the gradient field $\nabla_{g_0} \bar{G}$ can be written as*

$$\rho \nabla_{g_0} \bar{G} = -x_1 \frac{\partial}{\partial x_1} - x_2 \frac{\partial}{\partial x_2}$$

in the domain U .

Proof. Let us take local isothermal coordinates $u = (u_1, u_2)$ centered at \bar{y} , in which the metric reads as

$$g_0 = f(u) (du_1^2 + du_2^2)$$

for a positive function f . Therefore, Eq. (2.3) can be written in these coordinates as

$$\frac{\partial^2 \bar{G}}{\partial u_1^2} + \frac{\partial^2 \bar{G}}{\partial u_2^2} = -\delta_0,$$

so \bar{G} must be of the form

$$\bar{G} = -\frac{1}{2\pi} \log |u| + h(u),$$

with $|u|^2 := u_1^2 + u_2^2$ and h a harmonic function:

$$\frac{\partial^2 h}{\partial u_1^2} + \frac{\partial^2 h}{\partial u_2^2} = 0.$$

The gradient of \bar{G} can then be expressed as

$$\rho \nabla_{g_0} \bar{G} = - \left(u_1 - 2\pi |u|^2 \frac{\partial h}{\partial u_1} \right) \frac{\partial}{\partial u_1} - \left(u_2 - 2\pi |u|^2 \frac{\partial h}{\partial u_2} \right) \frac{\partial}{\partial u_2},$$

with $\rho(u) := 2\pi f(u) |u|^2$. The origin is then a hyperbolic zero of the vector field $\rho \nabla_{g_0} \bar{G}$ (which can be smoothly extended to the origin) and the corresponding eigenvalues are $(-1, -1)$. Hence Siegel's theorem ensures that $\rho \nabla_{g_0} \bar{G}$ can be linearized via a diffeomorphism that is an analytic function of the coordinates (u_1, u_2) and the claim follows. \square

In the following proposition we will characterize the structure of the trajectories of the field $\nabla_{g_0} \bar{G}$ in a neighborhood of a critical point of \bar{G} . In particular, we see that the trajectories approaching the critical point have a well-defined tangent. This proposition can be seen as a dynamical analog of Cheng's result on the critical points of eigenfunctions on surfaces [2]. Let us recall that the *stable* (resp. *unstable*) set of a zero z of a vector field is given by the points whose ω -limit (resp. α -limit) is exactly the point z .

Proposition 3.2. *Let z be a zero of the vector field $\nabla_{g_0} \bar{G}$ (possibly a removable singularity p_i) and let $m \geq 2$ be the degree of the lowest nonzero homogeneous term in the Taylor expansion of $\bar{G} - \bar{G}(z)$ at z (which is always finite). Then z is an isolated zero and the intersection of a small neighborhood of z in Σ with either its stable or unstable set is homeomorphic to the set*

$$\{\zeta \in \mathbb{C} : \zeta^m \in [0, 1]\},$$

In particular, the index of the point z is $1 - m$.

Proof. Let us take isothermal coordinates $u = (u_1, u_2)$ centered at z , in which the metric reads as

$$g_0 = f(u) (du_1^2 + du_2^2).$$

Therefore, \bar{G} is a harmonic function of u with respect to the flat metric:

$$(3.1) \quad \frac{\partial^2 \bar{G}}{\partial u_1^2} + \frac{\partial^2 \bar{G}}{\partial u_2^2} = 0.$$

It is therefore standard that \bar{G} is an analytic function of (u_1, u_2) .

Let h_m be the first nonzero homogenous polynomial that appears in the Taylor expansion of \bar{G} in these coordinates at 0, so that (with a slight abuse of notation)

$$(3.2a) \quad \bar{G}(u) - \bar{G}(0) = h_m(u) + O(|u|^{m+1}),$$

$$(3.2b) \quad f(u) \nabla_{g_0} \bar{G}(u) = \nabla h_m(u) + O(|u|^m),$$

Here we are denoting by ∇ the flat space gradient with respect to the coordinates u .

By Eq. (3.1), the homogeneous polynomial h_m is a harmonic function (with respect to the flat space Laplacian in the coordinates u), so it must be of the form

$$h_m = C \operatorname{Re} [e^{i\alpha} (u_1 + iu_2)^m]$$

for some real constants C, α . In particular, 0 is an isolated critical point of h_m , which readily implies that z is an isolated zero of $\nabla_{g_0} \bar{G}$.

Let us now consider polar coordinates $(r, \theta) \in (0, \epsilon) \times \mathbb{S}^1$ defined by $(u_1, u_2) = (r \cos \theta, r \sin \theta)$. In these coordinates one has

$$h_m(r, \theta) = C r^m \cos(m\theta + \alpha).$$

There is no loss of generality in setting $\alpha = 0$. We define the polar blow up of the gradient $\nabla_{g_0} \bar{G}$ at z using polar coordinates as the vector field

$$X := \frac{f}{Cm r^{m-2}} \nabla_{g_0} \bar{G} = \frac{1}{Cm r^{m-2}} (\nabla h_m + O(r^m)),$$

where we have used Eq. (3.2b). The blown-up trajectories are then given by

$$(3.3a) \quad \dot{r} = r \cos m\theta + O(r^2),$$

$$(3.3b) \quad \dot{\theta} = -\sin m\theta + O(r).$$

The blown-up critical points are thus $(0, \theta_k)$, with $\theta_k := k\pi/m$ and $k = 1, \dots, 2m$. The Jacobian matrix of X at $(0, \theta_k)$ is

$$(3.4) \quad DX(0, \theta_k) = \begin{pmatrix} (-1)^k & 0 \\ 0 & (-1)^{k+1} \end{pmatrix},$$

so these critical points are hyperbolic saddles. By blowing down, we immediately find that a deleted neighborhood of 0 consists exactly of $2m$ hyperbolic sectors of the vector field X .

Since the field X is proportional to the gradient field $\nabla_{g_0} \bar{G}$ through a factor that does not vanish but at z , this shows that the intersection with a small neighborhood of z with the its stable or unstable set is homeomorphic to

$$\{\zeta \in \mathbb{C} : \zeta^m \in [0, 1)\},$$

as claimed. Besides, the well known Bendixson formula for the index of a planar vector field asserts that the index of z is

$$\operatorname{ind}(z) = 1 - \frac{\text{number of hyperbolic sectors}}{2} = 1 - m,$$

as claimed. □

It is worth mentioning that the dynamics of the gradient of a harmonic function in a neighborhood of a critical point in dimension higher than 2 is much more involved, as is discussed in [9].

4. THE BASIN OF ATTRACTION OF THE POLE

In this section we will introduce the concept of basin of attraction associated with $\nabla_{g_0}\overline{G}$, as well as some convenient compactifications thereof. We shall see how this object and its boundary define a decomposition of the surface as the union of a disk and a 1-skeleton, as outlined in the Heuristic Principle in the Introduction. We shall also establish some properties of these sets.

The basin of attraction of the pole y is a key object in this paper, and can be thought of as the set of points of the surface M that approach y when flowed along the trajectories of the field $\nabla_g G$. However, in view of the characterization of M in terms of a compact surface Σ (the Uniformization Theorem 2.1), it is slightly more convenient to study the basin of attraction directly in this compact surface, so instead we will use the following

Definition 4.1. The *basin of attraction* is the set of points D in \mathcal{M} whose ω -limit along the trajectories of $\nabla_{g_0}\overline{G}$ is \bar{y} :

$$(4.1) \quad D := \{x \in \mathcal{M} : \omega(x) = \bar{y}\}.$$

Of course, by the relationship between $\nabla_{g_0}\overline{G}$ and $\nabla_g G$, the diffeomorphism $\Phi : \mathcal{M} \rightarrow M$ maps the basin D onto the set of points in M whose ω -limit along the integral curves of $\nabla_g G$ is the pole y . It is easy to prove that D is diffeomorphic to a disk.

It is clear that both D and its complement in \mathcal{M} are invariant sets under the local flow of $\nabla_{g_0}\overline{G}$. The complement,

$$\mathcal{F} := \mathcal{M} \setminus D,$$

will be a crucial object in the rest of the paper.

In the following proposition we shall prove a general result about sets of \mathcal{M} that are invariant under the local flow of $\nabla_{g_0}\overline{G}$ from which it stems an important property of \mathcal{F} as a corollary: that $\mathcal{M} \setminus D$ has empty interior and thus \mathcal{F} coincides with the boundary of the basin D in \mathcal{M} . Notice that, since the basin of attraction is contractible, \mathcal{F} cannot be empty unless the surface M is diffeomorphic to \mathbb{R}^2 .

Proposition 4.2. *Let $S \subset \mathcal{M}$ be an invariant set under the flow of $\nabla_{g_0}\overline{G}$ that does not contain the point \bar{y} and is relatively closed in \mathcal{M} . Then the interior of S is empty.*

Proof. In this proof, we will assume that we have enlarged the set \mathcal{M} and extended the function \overline{G} in the obvious way so that the removable singularities are points contained in \mathcal{M} and \overline{G} is also defined at these points. Let U denote a connected component of the interior of S . By Propositions 2.2 and 2.3, the disks D_i or the deleted points p_i that are not removable singularities behave as local minima of the function \overline{G} . Since $\Delta_{g_0}\overline{G} = 0$ both in U and in a neighborhood of any removable singularity p_i , the maximum principle for harmonic functions ensures that the maximum of \overline{G} must be attained at a point z of ∂U (possibly a removable singularity).

Since the boundary of U is invariant, z must be a critical point of \overline{G} , which is necessarily isolated by Proposition 3.2. As \overline{G} is smooth in a neighborhood of z , z

is an isolated maximum of \overline{G} in the closure \overline{U} and \overline{G} is increasing along the local flow ψ_t of $\nabla_{g_0}\overline{G}$, it follows that for any $\epsilon > 0$ there exists some $\delta > 0$ such that

$$\psi_t(B_{g_0}(z, \delta) \cap \overline{U}) \subset B_{g_0}(z, \epsilon)$$

for all $t > 0$ and

$$\bigcap_{t \geq 0} \psi_t(B_{g_0}(z, \delta) \cap \overline{U}) = \{z\}.$$

Hence there exists a region $B_{g_0}(z, \delta) \cap \overline{U}$ of nonzero measure whose ω -limit is z . This contradicts the fact that, since $\Delta_{g_0}\overline{G} = 0$ in a neighborhood of z , the local flow of $\nabla_{g_0}\overline{G}$ is area-preserving, so the set U must be empty. \square

Hence we immediately obtain the desired statement about \mathcal{F} :

Corollary 4.3. *The complement of D in \mathcal{M} coincides with the boundary of D in \mathcal{M} , that is,*

$$\mathcal{F} = \mathcal{M} \setminus D = \partial D.$$

Our goal now is to derive further properties of the set \mathcal{F} . For technical reasons, it is more convenient to consider the flow of an auxiliary vector field X defined in the whole compact surface Σ rather than that of $\nabla_{g_0}\overline{G}$, which is only defined on \mathcal{M} . For this, let us take a point q_i belonging to the interior of each disk D_i . Since the disk D_i retracts into q_i , one can take a diffeomorphism

$$\Psi : \mathcal{M} \rightarrow \Sigma \setminus \{p_1, \dots, p_{\lambda_1}, q_1, \dots, q_{\lambda_2}\}$$

which is equal to the identity outside a small neighborhood of the closed disks D_i (in particular, at \bar{y}).

Let us relabel the parabolic ends if necessary so that $\{p_i\}_{i=1}^{\lambda'_1}$ are the non-removable singularities, with $0 \leq \lambda'_1 \leq \lambda_1$. The auxiliary field is then defined as

$$X := F \Psi_* (\nabla_{g_0}\overline{G}),$$

where $F : \Sigma \rightarrow \mathbb{R}$ is a smooth nonnegative function that only vanishes at $\{\bar{y}\} \cup \mathcal{P}$, with

$$(4.2) \quad \mathcal{P} := \{p_1, \dots, p_{\lambda'_1}, q_1, \dots, q_{\lambda_2}\}$$

and is chosen so that X can be smoothly extended to the whole surface Σ . It is standard that such a factor always exists; notice, moreover, that we do not need to impose any additional conditions on this factor to ensure that X is also defined at the removable singularities (we only need to consider the obvious extension of \overline{G}). Throughout this paper, we will denote the flow of X by ϕ_t .

It is clear that the function $\overline{G} \circ \Psi^{-1}$ can then be extended to a continuous function $\widehat{G} : \Sigma \rightarrow [-\infty, +\infty]$ by setting

$$\widehat{G}(\bar{y}) := +\infty, \quad \widehat{G}(p_i) := -\infty, \quad \widehat{G}(q_j) := 0$$

for $1 \leq i \leq \lambda'_1$ and $1 \leq j \leq \lambda_2$. Hence it stems that the field X is gradient-like: the Lie derivative $\mathcal{L}_X \widehat{G}$ is strictly positive but at the zeros of the field X , which are the only points at which the function \widehat{G} can fail to be smooth. Moreover, the zeros of X are exactly the images under the diffeomorphism Ψ of the zeros of $\nabla_{g_0}\overline{G}$ (including those that correspond to removable singularities) and $\{y\} \cup \mathcal{P}$. Notice that the points in \mathcal{P} are precisely the global isolated minima of \widehat{G} : indeed,

by Proposition 2.3 if $\lambda_2 \geq 1$, then \overline{G} is positive and $\lambda'_1 = 0$. Besides, from this proposition it stems that the cardinality of \mathcal{P} is

$$(4.3) \quad \lambda' := \begin{cases} \lambda'_1 & \text{if } \lambda_2 = 0, \\ \lambda_2 & \text{if } \lambda_2 \geq 1. \end{cases}$$

The reason why we consider the field X is that, instead of directly analyzing the sets D and \mathcal{F} associated with the field $\nabla_{g_0}\overline{G}$, it is easier to consider the analogous dynamical objects for the field X . That is, we will consider the set \widehat{D} of the points in Σ whose ω -limit along the flow of X is \bar{y} , which we will still call the *basin of attraction* of the field X . Its associated *basin boundary* is then defined as

$$\widehat{\mathcal{F}} := \Sigma \setminus \widehat{D}.$$

Just as in the case of the set D , it is standard that \widehat{D} is diffeomorphic to a disk.

Remark 4.4. Proposition 4.2 and its proof are also valid, mutatis mutandis, for the field X . That is, if $S \subset \Sigma$ is a closed invariant set under the flow of X that does not contain the point \bar{y} , it has empty interior. In particular,

$$\widehat{\mathcal{F}} = \Sigma \setminus \widehat{D} = \partial \widehat{D}.$$

The following proposition completely characterizes the set $\widehat{\mathcal{F}}$ in terms of the zeros of the field X and their stable components. To state this result, let us introduce the notation

$$(4.4) \quad \mathcal{C} := \Psi[\{z \in \mathcal{M} \cup \{p_{\lambda'_1+1}, \dots, p_{\lambda_1}\} : \nabla_{g_0}\overline{G}(z) = 0\}]$$

for the image under the diffeomorphism Ψ of the critical points of \overline{G} , including those that may correspond to a removable singularity.

Proposition 4.5. *$\widehat{\mathcal{F}}$ is the union of the zeros of the field X other than \bar{y} and the stable sets of the zeros in the set \mathcal{C} :*

$$\widehat{\mathcal{F}} = \mathcal{P} \cup \bigcup_{z \in \mathcal{C}} W^s(z).$$

Furthermore, $\widehat{\mathcal{F}}$ is connected.

Proof. Remark 4.4 readily implies that $\widehat{\mathcal{F}}$ is connected and has empty interior. Let us now set

$$W := \mathcal{P} \cup \bigcup_{z \in \mathcal{C}} W^s(z).$$

By definition, it is clear that the ω -limit of any point $x \in W$ cannot be the point \bar{y} , so $W \subset \Sigma \setminus \widehat{D}$.

Let us now prove the converse implication: $\Sigma \setminus \widehat{D} \subset W$. For this, we shall show that the ω -limit of any $x \in \Sigma \setminus \widehat{D}$ must be a zero of the field X different from \bar{y} . We can obviously suppose that x is not a zero of X . Since X is gradient-like, $\phi_t x$ must tend to a zero z of X as $t \rightarrow \infty$ [15]. Besides, z must belong to the set \mathcal{C} because the points in \mathcal{P} are minima of \widehat{G} and the Lie derivative

$$\mathcal{L}_X \widehat{G}(\phi_t x) = \frac{d}{dt} \widehat{G}(\phi_t x)$$

is positive if x is not a zero of X . Hence we infer that $\widehat{\mathcal{F}} = W$. \square

It is clear that a good description of $\widehat{\mathcal{F}}$ immediately yields a complete characterization of the original set \mathcal{F} . In particular, Proposition 4.5 easily implies that, roughly speaking, \mathcal{F} consists of the set $\widehat{\mathcal{F}}$ and some additional segments that connect a point in $\widehat{\mathcal{F}}$ with a parabolic end of the surface. From this set, of course, we still have to remove the points in $\widehat{\mathcal{F}}$ corresponding to an end. This is the content of the following

Corollary 4.6. *The set \mathcal{F} is the image under the diffeomorphism Ψ^{-1} of the set*

$$\left(\widehat{\mathcal{F}} \cup \bigcup_{i=\lambda'_1+1}^{\lambda_1} \{ \phi_{-t} p_i : t > 0 \} \right) \setminus \{ p_1, \dots, p_{\lambda_1}, q_1, \dots, q_{\lambda_2} \}.$$

The closure of \mathcal{F} in Σ is connected.

Proof. It is clear that the image under Ψ^{-1} of the set $\widehat{\mathcal{F}}$ minus the ends

$$\{ p_1, \dots, p_{\lambda_1}, q_1, \dots, q_{\lambda_2} \}$$

must be contained in \mathcal{F} . Likewise, the image under Ψ^{-1} of any point x in \widehat{D} will be in the basin of attraction D unless at some positive time t the trajectory $\phi_t x$ passes through an end, which is necessarily a removable singularity since the other ends are all zeros of the field X . That is, for some $t > 0$ one would have

$$x = \phi_{-t} p_i \quad \text{for some } \lambda'_1 + 1 \leq i \leq \lambda_1.$$

This proves the formula in the statement. (Of course, among the removable singularities it would be enough to consider those that are not a zero of the field X .)

Therefore, the closure of \mathcal{F} in Σ is either empty (then $\widehat{\mathcal{F}}$ consists of a single point) or is diffeomorphic to the union of $\widehat{\mathcal{F}}$ and a finite number of segments whose endpoints are a removable singularity p_i as above and a zero of the field X (which obviously belongs to $\widehat{\mathcal{F}}$). Therefore, the closure of \mathcal{F} is connected. \square

5. STRUCTURE OF THE BASIN BOUNDARY AND BOUNDS FOR THE CRITICAL POINTS

In this section we will provide a full characterization of the basin boundary, thereby obtaining an upper bound for the number of critical points of the Green's function. The characterization of the basin boundary lay bare a strong connection between the dynamics of the field X (or, equivalently, $\nabla_g G$) and the topology and conformal structure of the surface.

The following theorem provides an upper bound for the number of critical points of \overline{G} (including those corresponding to removable singularities) in terms of the conformal properties of the surface, which appear through the number λ' introduced in Eq. (4.3). A straightforward consequence of this result is the purely topological bound presented in Theorem 1.1, which differs from the present statement in that here we are using additional information on the conformal structure of the surface Σ to sharpen the upper bound:

Theorem 5.1. *The number of zeros of the field $\nabla_{g_0} \overline{G}$ in*

$$\mathcal{M} \cup \{ p_{\lambda'_1+1}, \dots, p_{\lambda_1} \},$$

that is, the cardinal of the set \mathcal{C} , is not larger than $2\nu + \lambda' - 1$, and if this upper bound is attained then \overline{G} is Morse.

Before presenting the proof of this result, it is illustrative to sketch the argument in the easiest case: when M is diffeomorphic to \mathbb{R}^2 . In this case, the proof is similar to that of a result on the absence of critical points in some boundary value problems in the exterior of a bounded domain in \mathbb{R}^n that we proved in [5]. Of course, this particular case is elementary and could be easily treated using the uniformization theorem, but it serves to illustrate the basic dynamical ideas underlying the proof of the general situation.

Sketch of the proof when M is diffeomorphic to \mathbb{R}^2 . Let us analyze what happens in \mathcal{M} , which is diffeomorphic to the sphere minus a point p . Suppose that we have a zero z . By Proposition 3.2, its stable set (with the point z deleted) consists of at least two curves. The α -limit of each of these curves cannot be \bar{y} by Proposition 3.1, so as the vector field X is gradient-like either the curve approaches the end p or its α -limit is another zero z_1 . We can now apply the same argument for the zero z_1 , and if necessary to successive zeros z_2, z_3, \dots . Notice that $z_j \neq z_k$ for $j \neq k$, since otherwise we could have an invariant set (defined by the union of curves in the stable sets of zeros of the field) that is a Jordan curve not containing \bar{y} . Then this invariant set would separate the plane in two disjoint invariant sets with nonempty interior, contradicting Proposition 4.2.

Since the zeros are isolated in \mathcal{M} by Proposition 3.2, we can eventually take a union of curves in the stable sets of zeros z_j of the field (possibly infinitely many, but only accumulating at p) whose closure in the sphere is a Jordan curve that contains the point p but not \bar{y} . Again, this curve encloses an invariant set with nonempty interior that does not contain the point \bar{y} , in contradiction with Proposition 4.2. Hence there cannot be any zeros of $\nabla_{g_0} \overline{G}$ and the theorem follows. \square

In the general case, the proof is more involved and relies on a careful analysis of the saddle connections between zeros of the field. In the demonstration we need two lemmas that are presented right after the proof and make use of the same notation.

The proof is divided in two parts. First we show that the number of zeros of X is finite. Notice this is not trivial, since the zeros of X could accumulate at the set \mathcal{P} , which is associated with ends of the surface M , without contradicting the fact that critical points of \overline{G} are isolated by Proposition 3.2. However, we show that if there were an infinite number of critical points, there would be infinitely many closed invariant curves in Σ defining independent homology classes, which is forbidden by the fact that the fundamental group of Σ has finite rank. Roughly speaking, these closed invariant curves are constructed by successive continuation of the stable sets of some zeros of X . The second part of the proof consists in estimating the number of zeros of X using Hopf's index theorem and the characterization of the dynamics of X in a neighborhood of each zero.

Proof of Theorem 5.1. Let us start by recalling that the set \mathcal{P} , introduced in Eq. (4.2), consists of precisely λ' points, which are the minima of the function \widehat{G} . As X is gradient-like, the α - and ω -limit sets of any trajectory of this field are necessarily a zero of X . Moreover, the zeros of this field are obviously given by the set

$$\mathcal{Z} := \{\bar{y}\} \cup \mathcal{C} \cup \mathcal{P},$$

where we record here that the set \mathcal{C} , defined in (4.4), corresponds to the critical points of \overline{G} (possibly including removable singularities) under the diffeomorphism Ψ . The fact that the critical points of \overline{G} are isolated (by Proposition 3.2) guarantees that \mathcal{Z} does not accumulate but possibly at \mathcal{P} .

Let γ be a trajectory of the field X . This trajectory will be called *constant* if it consists of a single point. We will use the notation $\alpha(\gamma)$ and $\omega(\gamma)$ for the α - and ω -limit sets of γ . Let us introduce a partial order on the set of zeros \mathcal{Z} as follows. Given two points $x, x' \in \mathcal{Z}$, we shall write $x \succ x'$ if, for any open neighborhoods $U \ni x$ and $V \ni x'$ in Σ , there exist integers $p \leq 0$, $q \geq 0$ and nonconstant trajectories (of the field X) $\gamma_p, \dots, \gamma_q$ such that

- (i) $\omega(\gamma_p) \in U$, $\alpha(\gamma_q) \in V$,
- (ii) $\alpha(\gamma_j) = \omega(\gamma_{j+1})$ for $p \leq j \leq q - 1$.

Finiteness. We claim that the set of zeros \mathcal{Z} is finite. In order to prove this, let us assume the contrary. By Lemma 5.3 below, for each point $x \in \mathcal{C}$ there exists some point $p \in \mathcal{P}$ such that $x \succ p$. Since \mathcal{P} is finite, there exists some $p_1 \in \mathcal{P}$ such that one can choose a sequence $(x_k)_{k=1}^\infty$ of distinct points in \mathcal{C} with $x_k \succ p$. For each point x_k , Lemma 5.2 below yields a continuous path $\Gamma_{k,1} : [0, 1] \rightarrow \Sigma$ whose image is invariant under the field X and satisfies $\Gamma_{k,1}(0) = x_k$ and $\Gamma_{k,1}(1) = p_1$.

As a straightforward consequence of Proposition 3.2, the invariant set

$$W^s(x_k) \setminus \Gamma_{k,1}([0, 1])$$

is nonempty. If we let \tilde{x}_k be the α -limit of a trajectory contained in this set, we obviously have $x_k \succ \tilde{x}_k$. By Lemma 5.3, either $\tilde{x}_k \in \mathcal{P}$ or $\tilde{x}_k \succ p$ for some $p \in \mathcal{P}$. Since \mathcal{P} is finite, by Lemma 5.2 and possibly upon restricting ourselves to a subsequence that we still denote by $(x_k)_{k=1}^\infty$, we obtain a family of continuous paths $\Gamma_{k,2} : [0, 1] \rightarrow \Sigma$ whose image is invariant under X and such that $\Gamma_{k,2}(0) = x_k$ and $\Gamma_{k,2}(1) = p_2$ for some fixed $p_2 \in \mathcal{P}$ (possibly the same as p_1).

By construction, for each positive integer k the connected set

$$\Gamma_{k,1}([0, 1]) \cup \Gamma_{k,2}([0, 1]) \cup \Gamma_{k+1,1}([0, 1]) \cup \Gamma_{k+1,2}([0, 1])$$

consists of a continuous curve that connects the points x_k and x_{k+1} passing through p_1 and another continuous curve that connects the same pair of points x_k, x_{k+1} passing through p_2 . It is then evident that this set, which can have self-intersections, contains an invariant loop (continuous closed curve) Λ_k . One can obviously ensure that $\Lambda_k \neq \Lambda_{k'}$ for $k \neq k'$ (these loops can intersect, though), and that the point \bar{y} does not belong to any Λ_k .

For any integer j , the union of invariant loops

$$\bigcup_{k=1}^j \Lambda_k$$

cannot disconnect Σ , since a connected component of $\Sigma \setminus \bigcup_{k=1}^j \Lambda_k$ that does not contain \bar{y} would be an invariant set with nonempty interior, contradicting Proposition 4.2. Therefore, it is standard that the homology classes $[\Lambda_k] \in H_1(\Sigma; \mathbb{Z})$ defined by the cycle Λ_k must be independent for all $k = 1, 2, \dots$. This is impossible in a finitely generated surface, so we infer that the set \mathcal{C} is finite.

Upper bound. Let us now pass to bound the cardinality of the set \mathcal{C} . Suppose that the number λ_2 of removed disks is at least one, so that

$$\mathcal{P} = \{q_1, \dots, q_{\lambda_2}\}.$$

A first observation is that the points q_i are isolated zeros of the field X , by the finiteness of \mathcal{Z} , and are local repellers because they correspond to minima of the function \widehat{G} . Therefore, the index of X at these points is

$$\text{ind}(q_i) = 1.$$

Similarly, the point \bar{y} is an isolated zero which is a local attractor, so it has index 1. If we now apply Hopf's index theorem to the vector field X in Σ , we get that the sum of the indices of the zeros of X equals the Euler characteristic of Σ :

$$(5.1) \quad \text{ind}(\bar{y}) + \sum_{z \in \mathcal{C}} \text{ind}(z) + \sum_{i=1}^{\lambda_2} \text{ind}(q_i) = \chi(\Sigma) = 2 - 2\nu.$$

Since the index of each point $z \in \mathcal{C}$ is smaller than or equal to -1 by Proposition 3.2, plugging the values of the indices of \bar{y} and q_i we find

$$(5.2) \quad \#\mathcal{C} \leq - \sum_{z \in \mathcal{C}} \text{ind}(z) = 2\nu + \lambda_2 - 1.$$

The equality is not satisfied but perhaps when $\text{ind}(z) = -1$ for all $z \in \mathcal{C}$, that is, when \widehat{G} is Morse (by Proposition 3.2). This proves the theorem when $\lambda_2 \geq 1$.

Consider now the case where $\lambda_2 = 0$, so that

$$\mathcal{P} = \{p_1, \dots, p_{\lambda'_1}\}.$$

Arguing as before one easily finds that the index of X at each point p_i is

$$\text{ind}(p_i) = 1.$$

We now apply Hopf's index theorem to the vector field X in Σ to find

$$(5.3) \quad \text{ind}(\bar{y}) + \sum_{z \in \mathcal{C}} \text{ind}(z) + \sum_{i=1}^{\lambda'_1} \text{ind}(p_i) = \chi(\Sigma) = 2 - 2\nu,$$

so that the same argument as above yields

$$(5.4) \quad \#\mathcal{C} \leq - \sum_{z \in \mathcal{P}} \text{ind}(z) = 2\nu + \lambda'_1 - 1.$$

Again, the inequality being saturated at most when $\text{ind}(z) = -1$ for all $z \in \mathcal{C}$ (that is, when \widehat{G} is Morse). The theorem then follows. \square

Lemma 5.2. *Let $x, x' \in \mathcal{Z}$ such that $x \succ x'$. Then there exists a (not necessarily unique) injective continuous path $\Gamma : [0, 1] \rightarrow \Sigma$ such that:*

- (i) $\Gamma(0) = x$ and $\Gamma(1) = x'$.
- (ii) $\widehat{G} \circ \Gamma$ is strictly decreasing.
- (iii) The curve $\Gamma([0, 1])$ is invariant under the field X .

Proof. By the definition of the partial order \succ and Zorn's lemma, there exists a countable sequence $\{\gamma_j\}_{j=\bar{p}}^{\bar{q}}$ ($-\bar{p}, \bar{q} \in \mathbb{N} \cup \{\infty\}$) of nonconstant trajectories of the field X satisfying Conditions (i) and (ii) in the proof of Theorem 5.1, and such that

$$\lim_{j \rightarrow \bar{p}} \omega(\gamma_j) = x, \quad \lim_{j \rightarrow \bar{q}} \alpha(\gamma_j) = x'.$$

Let us consider any continuous parametrization

$$\Gamma : [0, 1] \rightarrow \overline{\bigcup_{j=\bar{p}}^{\bar{q}} \gamma_j(\mathbb{R})} \subset \Sigma$$

mapping 0 to x and 1 to x' . Since the Lie derivative $\mathcal{L}_X \hat{G}$ is positive in $\Sigma \setminus \mathcal{Z}$, \hat{G} is increasing along nonconstant trajectories, which implies that $\hat{G} \circ \Gamma$ is strictly decreasing (notice that the definition of Γ accounts for the fact that this function is decreasing instead of increasing). Therefore, Γ is injective. Moreover, the curve $\Gamma([0, 1])$ is clearly invariant because it is the union of trajectories. \square

Lemma 5.3. *For any $z \in \mathcal{C} \cup \{\bar{y}\}$ there exists some $p \in \mathcal{P}$ such that $z \succ p$. Moreover, there are no $x \in \mathcal{Z}$ such that $x \succ x$ or $p \succ x$ for some $p \in \mathcal{P}$ (i.e., the elements in \mathcal{P} are minimal with respect to the partial order).*

Proof. Let us take an element $z \in \mathcal{C}$. By Proposition 3.2, there exists a neighborhood U of z such that $(W^s(z) \cap U) \setminus \{z\}$ has at least two components C_1, C_2 and each C_i is a piece of a trajectory of X . Since X is gradient-like, the α -limit set of C_1 is another zero $z_1 \in \mathcal{C} \cup \mathcal{P}$.

If $z_1 \in \mathcal{P}$, the statement follows. Otherwise, we can repeat the previous argument replacing z by z_1 . As X is gradient-like, proceeding this way we obtain a sequence of distinct points z_1, z_2, \dots in \mathcal{Z} . If there is some point $z_k \in \mathcal{P}$, we are done, so we can assume that there is a sequence $(z_k)_{k=1}^\infty$ of distinct points of \mathcal{Z} with $z_k \succ z_{k+1}$. Since $\Psi^{-1}(\mathcal{C})$ consists of isolated points in \mathcal{M} by Proposition 3.2, it follows that $\text{dist}_{g_0}(z_k, \mathcal{P})$ tends to zero as $k \rightarrow \infty$. Hence there must exist some $p \in \mathcal{P}$ such that a subsequence $(z_{k'})_{k' \in J}$ tends to p . As above, it then follows that the initial sequence $(z_k)_{k=1}^\infty$ also tends to p because of the fact that p is an isolated minimum of \hat{G} and \hat{G} is increasing along trajectories.

Notice that the proof also applies when we start with the point \bar{y} instead of a point $z \in \mathcal{C}$. Finally, note that obviously $x \not\succ x$ for any $x \in \mathcal{Z}$ by Lemma 5.2, and that $p \not\succ x$ because each $p \in \mathcal{P}$ is an isolated minimum of \hat{G} and \hat{G} is increasing along the flow of the field X . \square

In the following theorem we show that the basin boundary $\hat{\mathcal{F}}$ encodes the topology of the surface Σ . In particular, the flow of the field X defines in a natural way a decomposition of Σ into a disk \hat{D} and its 1-skeleton $\hat{\mathcal{F}}$. This decomposition is similar to the one arising when one considers the cut locus of a point in Σ (with respect to some metric g_0) but it is generally different. Together with Proposition 4.5 and Corollary 4.6, this provides a rigorous reformulation of the Heuristic Principle stated in the Introduction.

Theorem 5.4. *The set $\hat{\mathcal{F}}$ is a connected graph with the same homology as the surface Σ :*

$$H_1(\hat{\mathcal{F}}; \mathbb{Z}) = \mathbb{Z}^{2\nu}.$$

Proof. We showed in Proposition 4.5 that $\hat{\mathcal{F}}$ consists of the zeros of X other than \bar{y} (that is, $\mathcal{C} \cup \mathcal{P}$) and their stable components, which are continuous curves with endpoints belonging to $\mathcal{C} \cup \mathcal{P}$. Since \mathcal{C} is finite by Theorem 5.1, it then follows that $\hat{\mathcal{F}}$ is a connected graph.

Let B be a small disk in Σ centered at \bar{y} and let us consider its image under the time- t flow of X , $\phi_t(B)$. It is apparent that $\Sigma \setminus \{\bar{y}\}$ deform retracts onto the set

$$\Sigma \setminus \phi_t(B),$$

for any $t \leq 0$. Moreover,

$$\widehat{\mathcal{F}} = \bigcap_{t \leq 0} (\Sigma \setminus \phi_t(B))$$

by the definition of the set $\widehat{\mathcal{F}}$. This ensures that $\widehat{\mathcal{F}}$ is a strong deformation retract of $\Sigma \setminus \{\bar{y}\}$, so it is well known that, $\widehat{\mathcal{F}}$ being a connected graph, we have the isomorphism of homology groups

$$H_1(\widehat{\mathcal{F}}; \mathbb{Z}) = H_1(\Sigma \setminus \{\bar{y}\}; \mathbb{Z}).$$

Since

$$H_1(\Sigma \setminus \{\bar{y}\}; \mathbb{Z}) = H_1(\Sigma \setminus B; \mathbb{Z}) = H_1(\Sigma; \mathbb{Z}) = \mathbb{Z}^{2\nu}$$

by a standard argument using the Mayer–Vietoris sequence, the theorem follows. \square

Remark 5.5. Because of the connection between \mathcal{F} and $\widehat{\mathcal{F}}$, Theorem 5.4 also gives very detailed information about the structure of the set \mathcal{F} . In particular, \mathcal{F} is a graph but is necessarily noncompact and possibly disconnected. Moreover, the rank of $H_1(\mathcal{F}; \mathbb{Z})$ is at most 2ν but can be strictly smaller than this number, as some of the cycles that appear in $\widehat{\mathcal{F}}$ can be killed after removing the points of Σ that correspond to the ends of the noncompact surface.

To conclude, let us present an illustrative example in which the different sets that we have been discussing in Sections 4 and 5 can be computed explicitly:

Example 5.6. Let M be the torus minus the point $p := (0, \pi)$, written in terms of the standard 2π -periodic coordinates on the torus. We choose a complete, conformally flat metric g on M , and fix the position of the pole at $y = (0, 0)$. It is then standard that Li and Tam’s procedure [11] gives rise to a Green’s function G invariant under the isometric transformations of the flat torus that fix the point y .

It is then clear that the curves

$$\{0\} \times \mathbb{S}^1, \quad \{\pi\} \times \mathbb{S}^1, \quad \mathbb{S}^1 \times \{0\}, \quad \mathbb{S}^1 \times \{\pi\}$$

are then invariant under the local flow of $\nabla_g G$. An easy argument using the symmetries then show that the points $(\pi, 0)$ and (π, π) are zeros of the field $\nabla_g G$, which must be nondegenerate (hence hyperbolic saddles). These are the only zeros because the upper bound in Theorem 5.1 is attained.

The set $\widehat{\mathcal{F}}$ consists of the two circles $\{\pi\} \times \mathbb{S}^1$ and $\mathbb{S}^1 \times \{\pi\}$, which respectively correspond to (two) saddle connections and to the trajectories of X that connect a saddle with the point p . Since $\widehat{\mathcal{F}}$ contains two independent cycles, $H_1(\widehat{\mathcal{F}}; \mathbb{Z})$ is isomorphic to the first homology group of the torus. The set \mathcal{F} is then given by

$$\mathcal{F} = \widehat{\mathcal{F}} \setminus \{p\},$$

so it consists of a closed curve and an open curve. In particular, $H_1(\mathcal{F}; \mathbb{Z}) = \mathbb{Z}$.

6. APPLICATIONS AND REMARKS

In this last section we will make some remarks about the connections between the critical points of the Green's function and the conformal properties of the surface. We will end with some comments about Green's functions on surfaces of infinite topological type.

Conformal structure and dynamics. A consequence of Theorem 5.1 is that by analyzing the dynamics of the gradient field $\nabla_g G$ (or, equivalently, of the field X) we can sometimes extract information about the conformal structure of the underlying surface. For example, a surfaces must satisfy very stringent geometric conditions in order to admit a Green's function without any critical points, as we show in the following

Proposition 6.1. *Let us suppose that the surface M has at least two ends. If all its ends are either hyperbolic or non-removable singularities, then any Green's function on M has at least one critical point.*

Proof. Eqs. (5.2) and (5.4) mean that, with the same notation as in the proof of Theorem 5.1,

$$(6.1) \quad - \sum_{z \in \mathcal{C}} \text{ind}(z) = 2\nu - 1 + \lambda'.$$

Let us call

$$\mathcal{C}_1 := \mathcal{C} \setminus \{p_{\lambda'_1+1}, \dots, p_{\lambda_1}\}$$

and $\mathcal{C}_2 := \mathcal{C} \setminus \mathcal{C}_1$. Any of these sets can be empty, and it is clear that the cardinality of \mathcal{C}_1 equals the number of critical points of the Green's function G in M . The cardinality of \mathcal{C}_2 is at most $\lambda_1 - \lambda'_1$.

As $\lambda' = \lambda'_1 + \lambda_2$, Eq. (6.1) can be rewritten as

$$(6.2) \quad - \sum_{z \in \mathcal{C}_1} \text{ind}(z) = 2\nu - 1 + \lambda'_1 + \lambda_2 + \sum_{z \in \mathcal{C}_2} \text{ind}(z).$$

If all the ends are hyperbolic, $\lambda_1 = \lambda'_1 = 0$, and if all ends are non-removable singularities, then $\lambda_1 = \lambda'_1$ and $\lambda_2 = 0$. Therefore, in both cases \mathcal{C}_2 is empty, so the RHS of (6.2) is nonzero and \mathcal{C}_1 cannot be empty. Since hyperbolic ends and non-removable singularities cannot coexist, the proposition is proved. \square

Surfaces of infinite topology. It is worth emphasizing that the case of surfaces whose fundamental group is not finitely generated is totally different from the case of surfaces of finite type. We shall next provide examples of surfaces that are not finitely generated both with an infinite number of critical points and without any critical points.

Example 6.2. Let us consider the case where M is a torus of infinite genus with two reflection symmetries (see Figure 1). For this, we can regard M as a surface embedded in \mathbb{R}^3 and invariant under the reflections

$$\Pi_1(x_1, x_2, x_3) := (-x_1, x_2, x_3) \quad \text{and} \quad \Pi_3(x_1, x_2, x_3) := (x_1, x_2, -x_3).$$

We will endow M with the metric g induced by the Euclidean metric in \mathbb{R}^3 and consider a Green's function G on M (satisfying Assumptions 1 and 2 in Section 2) with a pole at a point y invariant under the reflections Π_1, Π_3 . Li and Tam's

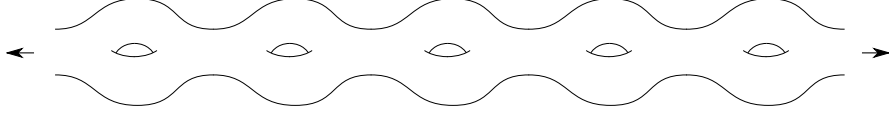


FIGURE 1. A torus of infinite genus with two reflection symmetries.

procedure can be used to obtain a Green's function with the same symmetries (that is

$$G = G \circ \Pi_j$$

for $j = 1, 3$).

It is straightforward that the intersection of the surface M with the plane $\{x_j = 0\}$ is invariant under the local flow of $\nabla_g G$, for $j = 1, 3$. A simple argument then shows that the intersection of M with the x_2 -axis must be invariant too. Since it consists of isolated points, these must therefore be either critical points of G or the pole y . These critical points are obviously infinite in number.

Example 6.3. Let us consider the unit disk

$$\mathbb{D}^2 := \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$$

with its hyperbolic metric g_1 . We let M be the unit disk with infinitely points removed as

$$M = \mathbb{D}^2 \setminus \left((0, 0) \cup \bigcup_{n=2}^{\infty} \left(0, \frac{1}{n} \right) \right).$$

The hyperbolic metric g_1 is not complete on M , but it is well known [18] that there is a conformally equivalent metric $g = \chi g_1$ such that (M, g) is complete.

Consider the minimal Green's function G with a pole at a point $y \in M$. By the conformal invariance of the Laplacian, this Green's function is precisely the minimal Green's function of the disk \mathbb{D}^2 with the hyperbolic metric g_1 (or rather its restriction to M). Since the gradient of the minimal Green's function of the hyperbolic disk (\mathbb{D}^2, g_1) does not vanish by Theorem 5.1, (M, g) is an example of a surface that is not finitely generated whose minimal Green's function does not have any critical points. Incidentally, notice that, although the proof of Theorem 5.1 does not apply to surfaces of infinite topological type, zero is exactly the upper bound one gets by recklessly applying the formula $2\nu - 1 + \lambda_2$ in this case (notice that for M we have $\nu = 0$, $\lambda_1 = \infty$, $\lambda'_1 = 0$, $\lambda_2 = 1$).

ACKNOWLEDGMENTS

This work is supported in part by the Spanish MINECO under grants FIS2011-22566 (A.E.), MTM2010-21186-C02-01 (D.P.-S.) and the ICMAT Severo Ochoa grant SEV-2011-0087 (A.E. and D.P.S.), and by Banco Santander-UCM under grant GR35/10-A-910556 (A.E.). The authors acknowledge the Spanish Ministry of Science for financial support through the Ramón y Cajal program.

REFERENCES

1. M. Brelot, G. Choquet, Espaces et lignes de Green. Ann. Inst. Fourier 3 (1951) 199–263.
2. S.Y. Cheng, Eigenfunctions and nodal sets, Comment. Math. Helv. 51 (1976) 43–55.

3. S.Y. Cheng, S.T. Yau, Differential equations on Riemannian manifolds and their geometric applications, *Comm. Pure Appl. Math.* 28 (1975) 333–354.
4. T.H. Colding, W.P. Minicozzi II, Large scale behavior of kernels of Schrödinger operators, *Amer. J. Math.* 119 (1997) 1355–1398.
5. A. Enciso, D. Peralta-Salas, Critical points and level sets in exterior boundary problems, *Indiana Univ. Math. J.* 58 (2009) 1947–1968.
6. A. Enciso, D. Peralta-Salas, Critical points of Green's functions on complete manifolds, *J. Differential Geom.* 92 (2012) 1–29.
7. A. Gabrielov, D. Novikov, B. Shapiro, Mystery of point charges. *Proc. Lond. Math. Soc.* 95 (2007) 443–472.
8. D. Gilbarg, J. Serrin, On isolated singularities of solutions of second order elliptic differential equations, *J. Anal. Math* 4 (1955/56) 309–340.
9. P. Goldstein, Gradient flow of a harmonic function in \mathbb{R}^3 , *J. Differential Equations* 247 (2009) 2517–2557.
10. A.A. Grigoryan, L. Saloff-Coste, Dirichlet heat kernel in the exterior of a compact set, *Comm. Pure Appl. Math.* 55 (2002) 93–133.
11. P. Li, L.F. Tam, Symmetric Green's functions on complete manifolds, *Amer. J. Math.* 109 (1987) 1129–1154.
12. P. Li, L.F. Tam, J. Wang, Sharp bounds for the Green's function and the heat kernel, *Math. Res. Lett.* 4 (1997) 589–602.
13. P. Li, S.T. Yau, On the parabolic kernel of the Schrödinger operator, *Acta Math.* 156 (1986) 153–201.
14. C.S. Lin, C.L. Wang, Elliptic functions, Green functions and the mean field equations on tori, *Ann. of Math.* 172 (2010) 911–954.
15. W. de Melo, J. Palis, *Geometric theory of dynamical systems*, Springer, New York, 1982.
16. M. Morse, Equilibrium points of harmonic potentials, *J. Anal. Math.* 23 (1970) 281–296.
17. M. Morse and S.S. Cairns, *Critical point theory in global analysis and differential topology*, Academic Press, New York, 1969.
18. K. Nomizu, H. Ozeki, The existence of complete Riemannian metrics, *Proc. Amer. Math. Soc.* 12 (1961) 889–891.
19. R. Shelton, Critical points of harmonic functions on domains in \mathbb{R}^3 , *Trans. Amer. Math. Soc.* 261 (1980) 137–158.
20. J.L. Walsh, *The location of critical points of analytic and harmonic functions*, AMS, New York, 1950.

INSTITUTO DE CIENCIAS MATEMÁTICAS, CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS,
28049 MADRID, SPAIN

E-mail address: `aenciso@icmat.es`

INSTITUTO DE CIENCIAS MATEMÁTICAS, CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS,
28049 MADRID, SPAIN

E-mail address: `dperalta@icmat.es`